

# Computing Multisensory Target Probabilities on a Neural Map

T. J. Anastasio<sup>1,2</sup>, P. E. Patton<sup>2</sup>

<sup>1,2</sup>Department of Molecular and Integrative Physiology and <sup>2</sup>Beckman Institute, University of Illinois at Urbana/Champaign, Urbana, IL, USA

**Abstract** - The superior colliculus is organized topographically as a neural map. The deep layers of the colliculus detect and localize targets in the environment by integrating input from multiple sensory systems. Some deep colliculus neurons receive input of only one sensory modality (unimodal) while others receive input of multiple modalities. Multimodal deep SC neurons exhibit multisensory enhancement, in which the response to input of one modality is augmented by input of another modality. Multisensory enhancement is magnitude dependent in that combinations of smaller single-modality responses produce larger amounts of enhancement. These findings are consistent with the hypothesis that deep colliculus neurons use sensory input to compute the probability that a target has appeared at their corresponding location in the environment. Multisensory enhancement and its magnitude dependence can be simulated using a model in which sensory inputs are random variables and target probability is computed using Bayes' Rule. Informational analysis of the model indicates that input of another modality can indeed increase the amount of target information received by a multimodal neuron, but only if input of the initial modality is ambiguous. Unimodal deep colliculus neurons may receive unambiguous input of one modality and have no need of input of another modality.

**Keywords** - Superior colliculus, computational neuroscience, multisensory integration, sensor fusion

## INTRODUCTION

Many neural structures are organized topographically, and neural maps can be found throughout the brain. One of the most well studied examples of a neural map is the superior colliculus (SC). The SC is located in the mammalian midbrain and is homologous to the optic tectum of non-mammals [1]. On grounds of differing connectivity and function, the SC can be divided into superficial and deep layers [2]. The superficial SC receives only visual input and sends it output primarily to the cortex. The deep SC integrates multisensory input and participates in the generation of orienting movements directed toward the source of sensory stimulation (target).

The deep SC in mammals receives convergent inputs from the visual, auditory, and somatosensory systems [3]. The deep SC sends its outputs to premotor circuits in the brainstem and spinal cord that control movements of the eyes and other structures. Neurons in the SC are organized topographically according to their receptive fields. Maps for the various sensory modalities are in register. The motor output of the SC is also topographically organized. Activation of neurons in a localized region of the SC leads to an orienting movement, such as a saccade of a stereotyped direction and magnitude.

Because the SC is organized topographically, it automatically localizes a stimulus source according to the corresponding location of the neurons in the SC map that are activated by the stimulus. Neurons in the deep SC can be activated by stimuli of more than one sensory modality [4-7]. Multimodal deep SC neurons can show the property of multisensory enhancement. Multisensory enhancement (MSE) is defined as the augmentation of the response of a neuron to a stimulus of one sensory modality by a stimulus of another modality. Percent MSE is computed as:

$$\%MSE = [(CM - SM_{\max}) / SM_{\max}] \times 100 \quad (1)$$

where CM is the combined-modality response and  $SM_{\max}$  is the larger of the two unimodal responses [6]. Percent MSE can range upwards of 1000%. Percent MSE is larger when the single-modality responses are smaller (Fig. 1). This property, known as inverse effectiveness, is perhaps the most revealing feature of MSE [3].

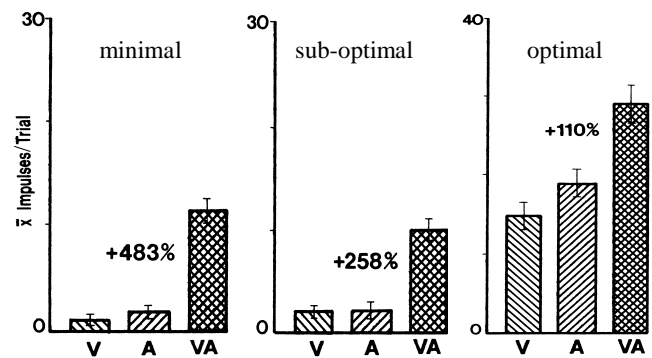


Fig. 1. Data from a single deep SC neuron illustrating the phenomena of multisensory enhancement and inverse effectiveness. (Data redrawn with permission from [3].)

The findings on MSE and inverse effectiveness show that multimodal deep SC neurons are performing a computation more complicated than simple summation. We have proposed a probabilistic model that can account for these findings [8]. The model provides a functional interpretation of MSE.

## PROBABILISTIC MODEL

We hypothesize that the response of a deep SC neuron to multisensory input is proportional to the conditional probability that a target has appeared in its receptive field. A schematic diagram illustrating this hypothesis is shown in Fig. 2. The target is represented as binary random variable  $T$  where  $T=1$  when the target is present and  $T=0$  when it is absent. As an example, we arbitrarily let the target-present and target-absent prior probabilities be  $P(T=1) = 0.1$  and

## Report Documentation Page

<b>Report Date</b> 25 Oct 2001	<b>Report Type</b> N/A	<b>Dates Covered (from... to)</b> -
<b>Title and Subtitle</b> Computing Multisensory Target Probabilities on a Neural Map		<b>Contract Number</b>
		<b>Grant Number</b>
		<b>Program Element Number</b>
<b>Author(s)</b>	<b>Project Number</b>	
	<b>Task Number</b>	
	<b>Work Unit Number</b>	
<b>Performing Organization Name(s) and Address(es)</b> Department of Molecular and Integrative Physiology and Beckman Institute Urbana, IL		<b>Performing Organization Report Number</b>
<b>Sponsoring/Monitoring Agency Name(s) and Address(es)</b> US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500		<b>Sponsor/Monitor's Acronym(s)</b>
		<b>Sponsor/Monitor's Report Number(s)</b>
<b>Distribution/Availability Statement</b> Approved for public release, distribution unlimited		
<b>Supplementary Notes</b> Papers from the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.		
<b>Abstract</b>		
<b>Subject Terms</b>		
<b>Report Classification</b> unclassified	<b>Classification of this page</b> unclassified	
<b>Classification of Abstract</b> unclassified	<b>Limitation of Abstract</b> UU	
<b>Number of Pages</b> 4		

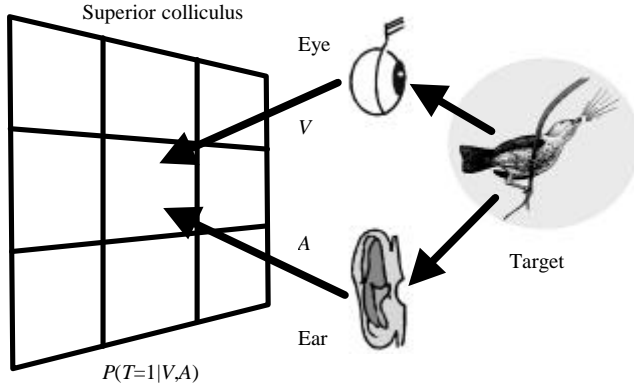


Fig. 2. Schematic diagram of the Bayes' Rule model of multisensory enhancement in the superior colliculus. (Redrawn with permission from [8].)

$P(T=0) = 0.9$ , respectively. The inputs to the deep SC, being neural, are modeled as stochastic. Discrete random variables  $V$  and  $A$  represent the number of spikes per unit time (0.25 s) in the inputs from the visual and auditory systems. The likelihoods of  $V$  or  $A$  represent the probability of observing different levels of input  $V$  or  $A$ , given that the sensory systems are being driven by a target ( $T=1$ ) or are only spontaneously active ( $T=0$ ). We model  $P(V|T=1)$ ,  $P(V|T=0)$ ,  $P(A|T=1)$ , or  $P(A|T=0)$  using Poisson distributions where the driven means are larger than the spontaneous means (Fig. 3). For simplicity, we also assume that  $V$  and  $A$  are conditionally independent given the target. Because  $P(T)$  and  $P(V|T)$  are properties of the sensory systems and the environment, we assume that they can be represented by the nervous system.

Given these distributions, the conditional probability of the target given the input can be computed using Bayes' Rule [9,10]. For example, consider a deep SC neuron receiving only visual input  $V$ . The unimodal Bayesian probability of the target is:

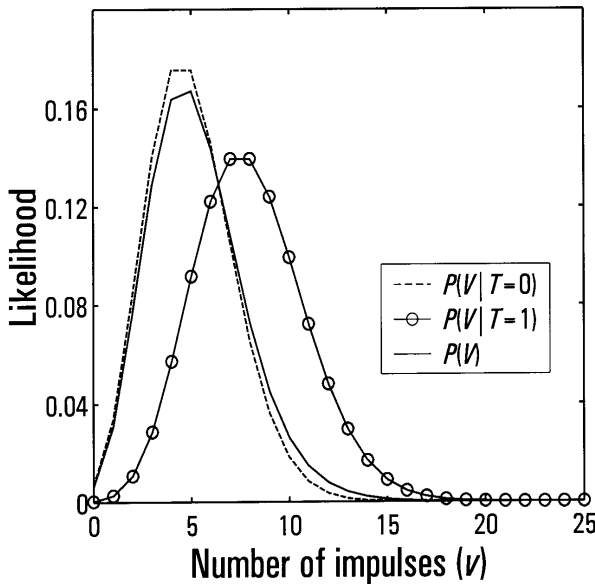


Fig. 3. Spontaneous and driven input likelihoods modeled as Poisson distributions having different means. (Redrawn with permission from [8].)

$$P(T=1|V) = P(V|T=1) P(T=1) / P(V). \quad (2)$$

By the principle of total probability [9,10]:

$$P(V) = P(V|T=1) P(T=1) + P(V|T=0) P(T=0). \quad (3)$$

The bimodal Bayesian target probability with  $V$  and  $A$  is:

$$P(T=1|V,A) = P(V,A|T=1) P(T=1) / P(V,A). \quad (4)$$

Because we assume, for simplicity, that the sensory inputs are conditionally independent given the target:

$$P(V,A|T=1) = P(V|T=1) P(A|T=1), \quad (5)$$

and again by the principle of total probability:

$$P(V,A) = P(V|T=1) P(A|T=1) P(T=1) + P(V|T=0) P(A|T=0) P(T=0). \quad (6)$$

Bayes' Rule in the bimodal case can be used to simulate MSE. To do this, we evaluate  $P(T=1|V,A)$  when both inputs  $V$  and  $A$  are driven (and  $V=A$  for simplicity), or when  $V$  or  $A$  is allowed to vary while the other is fixed at the spontaneous mean. The driven means for this example are 10 for  $V$  and 8 for  $A$ , and the spontaneous means are 5 for both. As  $V$  and/or  $A$  increase, we find that  $P(T=1|V,A)$  transitions sigmoidally from 0 to 1. The curve begins rising for smaller input values when both  $V$  and  $A$  are driven than when one is driven and the other is spontaneous. The simulated response in the both-driven case [ $P(T=1|V,A)$ ], and in either of the driven/spontaneous cases [ $P(T=1|V,(A=5))$  or  $P(T=1|(V=5),A)$ ] are plotted in Fig. 4. The curves show that the Bayesian probability of the target, especially for small inputs, can be many times higher in the both-driven case than in the driven/spontaneous cases.

This modeling result on MSE is used to simulate inverse effectiveness. Values of  $V$  and  $A$  are chosen that have similar driven/spontaneous Bayesian probabilities at three levels (minimal, sub-optimal, and optimal), and the both-driven Bayesian probability is computed using the same driven input values at each level. Percent MSE is computed using (1) where the both-driven and the larger driven/spontaneous probabilities are substituted for CM and

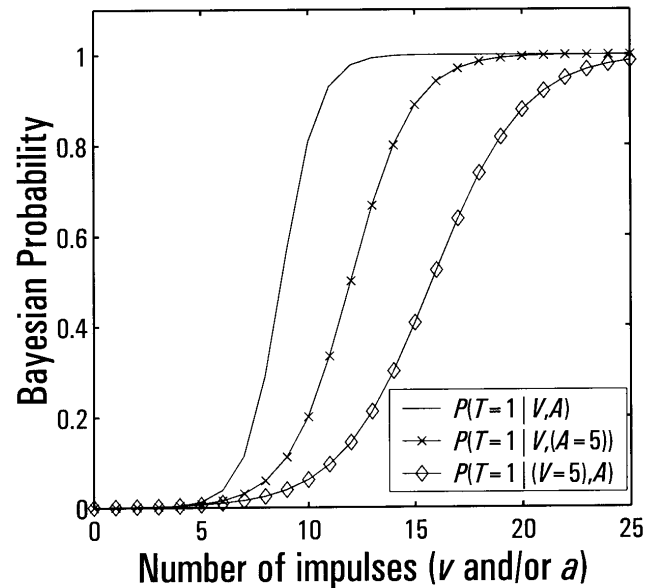


Fig. 4. Bimodal Bayesian target probabilities in the both-driven and the driven-spontaneous cases. (Redrawn with permission from [8].)

$SM_{max}$ , respectively. The Bayesian probabilities and their %MSE values are shown in Fig. 5. The probability values chosen illustrate the broad range of %MSE that can be simulated using the Bayesian model. The simulated results capture the essence of inverse effectiveness as it has been observed experimentally (Fig. 1).

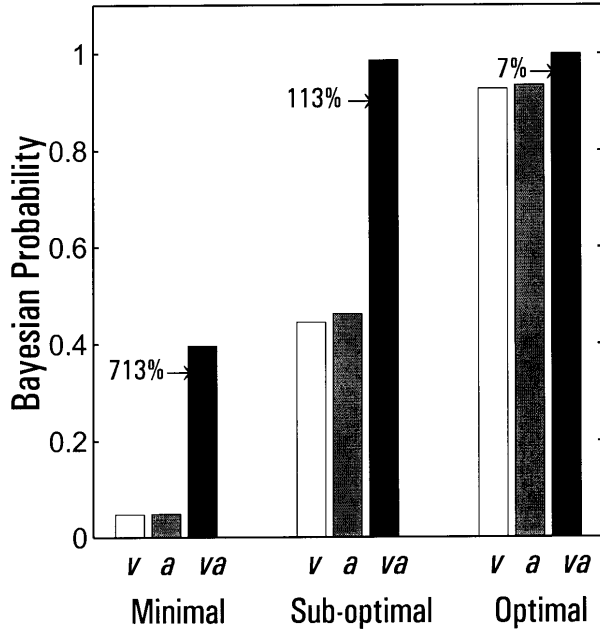


Fig. 5. Simulating multisensory enhancement and inverse effectiveness using the Bayes' Rule model. (Redrawn with permission from [8].)

Interpreting the responses of SC neurons as proportional to the probability of a target provides a functional explanation for inverse effectiveness. A large unimodal stimulus provides overwhelming evidence of a target. The Bayesian probability of a target would be close to 1 and a stimulus of another modality would not increase it much. However, if a unimodal stimulus is small, the Bayesian probability of a target can be 0 or near 0. Integrating a stimulus of another modality can dramatically increase the probability that a target has appeared. The qualitative correspondence between the Bayesian probabilities and the multisensory responses of deep SC neurons strongly supports the hypothesis that deep SC neurons compute the probability of a target given their multisensory inputs. The Bayes' Rule model of MSE serves as the basis for further analysis of multisensory integration.

#### INFORMATIONAL ANALYSIS

MSE may improve the ability of deep SC neurons to detect targets, yet many deep SC neurons are not multisensory. Respectively in cat and monkey, 46% and 73% of deep SC neurons are unimodal, while fewer than 50% are bimodal and fewer than 10% are trimodal in either species [11]. It is not obvious why so many deep SC neurons are unimodal despite the apparent usefulness of multisensory input. MSE

may be a mechanism whereby the deep SC decreases the uncertainty associated with some unimodal inputs. Measuring uncertainty is the domain of information theory. Information theoretic analysis offers a possible explanation for the puzzling finding that some deep SC neurons are multimodal while others are not [13].

The self-information, or information content of an event such as the presence of the target  $T$ , is given by [10,12]:

$$I(T=1) = -\log_2(P(T=1)) \quad (7)$$

where  $\log_2$  is the logarithm to the base two. Thus, the information content of the event, measured in units of bits, goes up as its probability goes down. Since the binary target is a random variable,  $I(T)$  is also a random variable. The expected value of this information is called the entropy and is denoted by  $H(T)$ :

$$H(T) = -[P(T=1) \log_2(P(T=1)) + P(T=0) \log_2(P(T=0))] \quad (8)$$

where  $x \log_2(x)$  is defined to be 0 when  $x=0$ . Entropy is a measure of the uncertainty associated with a random variable. For our prior distribution of  $T$  [ $P(T=1) = 0.1$  and  $P(T=0) = 0.9$ ] entropy  $H(T) = 0.469$  bits. The conditional entropy is the uncertainty that remains about the target after sensory input has been taken into account. For the unimodal and bimodal cases of MSE the conditional entropy would be computed respectively as:

$$H(T|V) = -\sum P(V) \sum P(T|V) \log_2(P(T|V)) \text{ and}$$

$$H(T|V,A) = -\sum \sum P(V,A) \sum P(T|V,A) \log_2(P(T|V,A)) \quad (9)$$

where the inner summations are over the values of the target and the outer summations are over the values of the sensory inputs. These formulae measure the uncertainty associated with  $T$  given inputs  $V$ , or  $V$  and  $A$ . Conditioning reduces entropy, so  $H(T|V,A) \leq H(T|V)$ .

The amount by which unimodal input  $V$ , or bimodal inputs  $V$  and  $A$ , reduces the uncertainty associated with  $T$  is equivalent to the average information gain:

$$I(T;V) = H(T) - H(T|V) \text{ and}$$

$$I(T;V,A) = H(T) - H(T|V,A). \quad (10)$$

The average information gain is equivalent to the mutual information between the target and inputs of one or two modalities. Mutual information can be used to quantify the ability of inputs having different statistical structures to provide target information.

The inputs themselves can be characterized using another information theoretic measure known as relative entropy, or the Kulback-Leibler divergence  $D$ . We assume that the ability of inputs  $V$  and  $A$  to provide information to deep SC neurons about the target is related to the divergence between their spontaneous and driven likelihood distributions. For the unimodal and bimodal cases, this divergence can be quantified as follows:

$$D(P(V|T=0)||P(V|T=1)) = \sum P(V|T=0)$$

$$\log_2(P(V|T=0)/P(V|T=1)) \text{ and}$$

$$D(P(V,A|T=0)||P(V,A|T=1)) = \sum \sum P(V,A|T=0)$$

$$\log_2(P(V,A|T=0)/P(V,A|T=1)) \quad (11)$$

where the summations are over the values of the sensory inputs. The divergence between spontaneous and driven likelihoods were computed for Poisson distributed inputs  $V$ ,

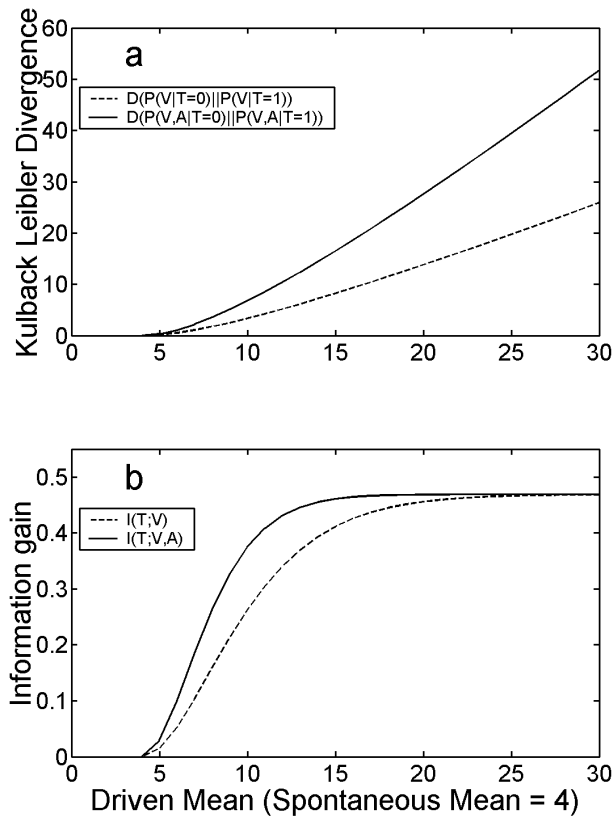


Fig. 6. Kulback Leibler divergence and information gain for Poisson distributed likelihoods as the driven mean is increased from 4 to 30 while the spontaneous mean is fixed at 4 impulses per unit time. (Redrawn with permission from [13].)

or  $V$  and  $A$ . The spontaneous means were fixed at 4 for both  $V$  and  $A$ , and their driven means were increased together from 4 to 30. The computed divergences  $D$  are shown in Fig. 6a. Consistent with intuition, the divergence  $D$  increases as the difference between the spontaneous and driven means increases. Also, the divergence increases faster for  $V$  and  $A$  than for  $V$  alone. It might also be consistent with intuition that the information provided by the inputs should increase as the difference between their spontaneous and driven likelihoods increases, and that bimodal inputs should always provide more information about the target than unimodal inputs. Such is not the case.

The mutual information, or information gained about  $T$  from inputs  $V$ , or  $V$  and  $A$ , is computed for the inputs as the driven means increase, and the results are plotted in Fig. 6b. The figure shows that information gain does increase as the driven means increase, and that it increases faster for bimodal than for unimodal inputs. However, both curves plateau at the same level [ $H(T) = 0.469$ ]. Thus, as the divergence between the spontaneous and driven likelihoods increases, the input reaches a point where it provides complete information about the target and cannot provide more. For the same reason, after the divergence between the spontaneous and driven likelihoods reaches the point where the information they provide has plateaued, a bimodal

input provides no more information than a unimodal input. This information theoretic analysis may have important implications for understanding MSE and the difference between unimodal and bimodal deep SC neurons.

As Fig. 6b shows, an input with a small divergence between the spontaneous and driven likelihoods carries relatively little information about the target, but this amount can be increased by an input of another modality (Fig. 6b for driven means less than about 15). In contrast, an input with a large divergence between the spontaneous and driven likelihoods carries the maximum amount of information about the target, and input of another modality is superfluous (Fig. 6b for driven means greater than about 25). Thus, bimodal deep SC neurons may be those that receive more uncertain unimodal input and require input of another modality to increase the amount of information that is transmitted to them. Unimodal deep SC neurons may be those that receive input of one modality that already supplies the maximum amount of information.

It would be difficult to test this theory directly because it is practically impossible to characterize the statistical structure of the inputs to real SC neurons. However, information theoretic ideas can be used as the basis for computational models of the development of MSE, and these can be tested against neurophysiological data from the deep SC of developing animals.

## REFERENCES

- [1] H. Vanegas, Ed. *Comparative Neurology of the Optic Tectum*, New York: Plenum Press, 1984.
- [2] R.L. Wurtz, and M.E. Goldberg, Eds. *The Neurobiology of Saccadic Eye Movements*, Amsterdam: Elsevier, 1989.
- [3] B.E. Stein, and M.A. Meredith, *The Merging of the Senses*, The MIT Press: Cambridge, 1993.
- [4] M.A. Meredith, and B.E. Stein, "Interactions among converging sensory inputs in the superior colliculus," *Science*, vol. 221(4608), pp. 389-391, 1983.
- [5] M.A. Meredith, and B.E. Stein, "Spatial factors determine the activity of multisensory neurons in cat superior colliculus," *Brain Res.*, vol. 365(2), pp. 350-354, 1986.
- [6] M.A. Meredith, and B.E. Stein, "Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration," *J. Neurophysiol.*, vol. 56(3), pp. 640-662, 1986.
- [7] A.J. King, and A.R. Palmer, "Integration of visual and auditory information in bimodal neurons in guinea-pig superior colliculus," *Exp Brain Res.*, vol. 60, pp. 492-500, 1985.
- [8] T.J. Anastasio, P.E. Patton, and K. Belkacem-Boussaid, "Using Bayes' rule to model multisensory enhancement in the superior colliculus," *Neural Computation*, vol. 12, pp. 997-1019, 2000.
- [9] C.W. Helstrom, *Probability and Stochastic Processes for Engineers*, Macmillan: New York, 1984.
- [10] D. Applebaum, *Probability and Information: An Integrated Approach*, Cambridge University Press: Cambridge, 1996.
- [11] M.T. Wallace, and B.E. Stein, "Sensory organization of the superior colliculus in cat and monkey," *Prog. Brain Res.*, vol. 112, pp. 301-311, 1996.
- [12] T.M. Cover, and J.A. Thomas, *Elements of Information Theory*, Wiley: New York, 1991.
- [13] P.E. Patton, K. Belkacem-Boussaid, and T.J. Anastasio, "Multimodality in the superior colliculus: an information theoretic analysis," *Cog. Brain Res.*, in press.